

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently amended) A method for assigning traffic buckets to a cache system, the method comprising:

a) when a new cache system starts up in a cache cluster having a plurality of cache systems among which a plurality of total buckets are to be allocated, determining a full bucket allocation for the new cache system and assigning a portion of the full bucket allocation to the new cache system that was determined at start up;

b) periodically determining a load of the new cache system;

c) each time it is periodically determined that the new cache system is underloaded and buckets have not been previously shed from the new cache system, slowly assigning a portion of the full bucket allocation that was determined at startup unless the cache cluster is operating at a maximum load;

d) when buckets have not been previously shed from the new cache system and each time it is periodically determined that the new cache system is underloaded and buckets have been shed previously from the new cache system, slowly re-assigning ~~or~~ a portion of previously shed the buckets that have been previously shed from the new cache system to the new cache system ~~when buckets have been previously shed from the new cache system,~~ unless the cache cluster is operating at a maximum load; and

ed) each time it is periodically determined that the new cache system is overloaded, shedding a portion of the buckets previously assigned to the new cache system,

wherein each bucket portion corresponds to a portion of the total traffic being handled by the cache cluster.

2. (original) A method as recited in claim 1 further comprising assigning the full bucket allocation to the new cache system when the cache cluster is operating at a maximum load.

3. (cancelled)

4. (previously presented) A method as recited in claim 1, wherein shedding a portion of the buckets previously assigned to the new cache comprises:

when no buckets have been previously shed, periodically shedding a portion of the assigned buckets from the new cache system;

when buckets have been previously shed, periodically shedding a portion of a number of buckets that were previously shed from the new cache system;

5. (currently amended) A method as recited in claim 4 further comprising after the full allocation of buckets is assigned to the new cache system or the last assigned or shed number of buckets for the new cache system is less than or equal to one:

halting operations (b) through (ed);

e) periodically monitoring a load of each of the other cache systems, including the new cache system, within the cache cluster;

f) when any of the cache systems is overloaded, shedding ~~the~~ a minimum number of buckets from the overloaded cache system; and

g) when any of the cache systems is underloaded, adding the minimum number of buckets to the underloaded cache system.

6. (currently amended) A method as recited in claim 5, wherein the minimum number of buckets equals a single bucket, the portion of the ~~unassigned-buckets~~ full bucket allocation equals a half of the ~~unassigned-buckets~~ full bucket allocation, and the portion of a number of buckets that were previously shed equals half of the number of buckets that were previously shed.

7. (currently amended) A method as recited in claim 1 wherein the operations (b) through (e) are performed until the full allocation has been assigned to the new cache system or a minimum number of buckets have been added to or shed from the new cache system.

8. (currently amended) A method as recited in claim 7 wherein operations (b) through (e) are performed about every 30 seconds.

9. (Cancelled)

10. (Cancelled)

11. (previously presented) A method as recited in claim 5 further comprising:

when an existing cache system leaves the cache cluster or shuts down,
determining a new bucket allocation for each of the remaining cache systems; and
assigning buckets to the remaining cache system using the new bucket allocation with operations (e) through (g).

12. (original) A method as recited in claim 1, wherein the full bucket allocation is equal to a number of buckets allocated to each existing cache system within the cache cluster.

13. (original) A method as recited in claim 1, wherein the full bucket allocation is not equal to a number of buckets allocated to each existing cache system within the cache cluster.

14. (original) A method as recited in claim 13, further comprising receiving a weight value from the new cache system indicating a percentage of the total buckets to allocate the new cache system.

15. (previously presented) A method as recited in claim 4 further comprising:
receiving load information from the new cache, the load information indicating whether the new cache system is overloaded; and
using the load information to determine whether the new cache is overloaded.

16. (original) A method as recited in claim 15 wherein load information is periodically received from the new cache system.

17. (currently amended) A method as recited in claim 15 wherein the load information further indicates a number of buckets to shed or add, wherein the portion of the number of buckets that were previously shed from the new cache system, ~~the portion of the unassigned buckets,~~ and the portion of the assigned buckets are equal to the indicated number of buckets to shed or add.

18. (Currently amended) A computer system operable to assign traffic buckets to a cache system, comprising:

a memory; and

a processor coupled to the memory,

a) when a new cache system starts up in a cache cluster having a plurality of cache systems among which a plurality of total buckets are to be allocated, determining a full bucket allocation for the new cache system

and assigning a portion of the full bucket allocation to the new cache system that was determined at start up;

b) periodically determining a load of the new cache system;

c) each time it is periodically determined that the new cache system is underloaded and buckets have not been previously shed from the new cache system, slowly assigning a portion of the full bucket allocation that was determined at startup unless the cache cluster is operating at a maximum load;

~~d) when buckets have not been previously shed from the new cache system and~~ each time it is periodically determined that the new cache system is underloaded and buckets have been shed previously from the new cache system, slowly re-assigning ~~er~~ a portion of ~~previously shed~~ the buckets that have been previously shed from the new cache system to the new cache system ~~when buckets have been previously shed from the new cache system~~, unless the cache cluster is operating at a maximum load; and

~~ed~~) each time it is periodically determined that the new cache system is overloaded, shedding a portion of the buckets previously assigned to the new cache system,

wherein each bucket portion corresponds to a portion of the total traffic being handled by the cache cluster.

19. (original) A computer system as recited in claim 18, wherein at least one of the memory and the processor are further adapted to provide:

assigning the full bucket allocation to the new cache system when the cache cluster is operating at a maximum load..

Application No.: 09/608,549
CISCP151/3223

20. (currently amended) A computer system as recited in claim 18, wherein at least one of the memory and the processor are further adapted to provide after the full allocation of buckets is assigned to the new cache system or the last assigned or shed number of buckets for the new cache system is less than or equal to one:

halting operations (b) through (ed);

e) periodically monitoring a load of each of the other cache systems, including the new cache system, within the cache cluster;

f) when any of the cache systems is overloaded, shedding ~~the~~ a minimum number of buckets from the overloaded cache system; and

g) when any of the cache systems is underloaded, adding the minimum number of buckets to the underloaded cache system.

21. (Cancelled)

22. (currently amended) A computer system as recited in claim ~~21~~ 18, wherein the portion of the number of buckets that were previously shed from the new cache system, the portion of the ~~unassigned buckets~~ full bucket allocation, and the portion of the assigned buckets are equal to a half portion.

23. (cancelled)

24. (currently amended) A computer system as recited in claim ~~18~~ 20 wherein at least one of the memory and the processor are further adapted to provide:

when an existing cache system leaves the cache cluster or shuts down, determining a new bucket allocation for each of the remaining cache systems; and assigning buckets to the remaining cache system using the new bucket allocation with operations (e) through (g).

25. (Currently amended) A computer program product for assigning traffic buckets to a cache system, the computer program product comprising:

at least one computer readable medium;

computer program instructions stored within the at least one computer readable product configured to:

a) when a new cache system starts up in a cache cluster having a plurality of cache systems among which a plurality of total buckets are to be allocated, determining a full bucket allocation for the new cache system and assigning a portion of the full bucket allocation to the new cache system that was determined at start up;

b) periodically determining a load of the new cache system;

c) each time it is periodically determined that the new cache system is underloaded and buckets have not been previously shed from the new cache system, slowly assigning a portion of the full bucket allocation that was determined at startup unless the cache cluster is operating at a maximum load;

~~d) when buckets have not been previously shed from the new cache system and~~ each time it is periodically determined that the new cache system is underloaded and buckets have been shed previously from the new cache system, slowly re-assigning ~~or~~ a portion of ~~previously shed~~ the buckets that have been previously shed from the new cache system to the new cache system ~~when buckets have been previously shed from the new cache system,~~ unless the cache cluster is operating at a maximum load; and

~~ed)~~ each time it is periodically determined that the new cache system is overloaded, shedding a portion of the buckets previously assigned to the new cache system,

Application No.: 09/608,549
CISCP151/3223

wherein each bucket portion corresponds to a portion of the total traffic being handled by the cache cluster.

26. (previously presented) A computer program product as recited in claim 25, wherein the computer program instructions are further configured to assign the full bucket allocation to the new cache system when the cache cluster is operating at a maximum load..

27. (cancelled)

28. (previously presented) A computer program product as recited in claim 25, wherein shedding a portion of the buckets previously assigned to the new cache comprises:

when no buckets have been previously shed, periodically shedding a portion of the assigned buckets from the new cache system;

when buckets have been previously shed, periodically shedding a portion of a number of buckets that were previously shed from the new cache system.

29. (currently amended) A computer program product as recited in claim 28 wherein the computer program instructions are further configured to after the full allocation of buckets is assigned to the new cache system or the last assigned or shed number of buckets for the new cache system is less than or equal to one:

halt operations (b) through (de);

e) periodically monitor a load of each of the other cache systems, including the new cache system, within the cache cluster;

f) when any of the cache systems is overloaded, shed the a minimum number of buckets from the overloaded cache system; and

g) when any of the cache systems is underloaded, add the minimum number of buckets to the underloaded cache system.

30. (currently amended) A computer program product as recited in claim 29, wherein the minimum number of buckets equals a single bucket, the portion of the ~~unassigned buckets~~ full bucket allocation equals a half of the ~~unassigned buckets~~ full bucket allocation, and the portion of a number of buckets that were previously shed equals half of the number of buckets that were previously shed.

31. (currently amended) A computer program product as recited in claim 25, wherein the operations (b) through (de) are performed until the full allocation has been assigned to the new cache system or a minimum number of buckets have been added to or shed from the new cache system

32. (currently amended) A computer program product as recited in claim 31, wherein operations (b) through (de) are performed about every 30 seconds.

33. (cancelled)

34. (previously presented) A computer program product as recited in claim 29, wherein the computer program instructions are further configured to:

when an existing cache system leaves the cache cluster or shuts down,
determine a new bucket allocation for each of the remaining cache systems; and
assign buckets to the remaining cache system using the new bucket allocation with operations (e) through (g).

35. (previously presented) A computer program product as recited in claim 25, wherein the full bucket allocation is equal to a number of buckets allocated to each existing cache system within the cache cluster.

36. (previously presented) A computer program product as recited in claim 25, wherein the full bucket allocation is not equal to a number of buckets allocated to each existing cache system within the cache cluster.

37. (previously presented) A computer program product as recited in claim 36, wherein the computer program instructions are further configured to receive a weight value from the new cache system indicating a percentage of the total buckets to allocate the new cache system.

38. (previously presented) A computer program product as recited in claim 28, wherein the computer program instructions are further configured to:

receive load information from the new cache, the load information indicating whether the new cache system is overloaded; and

use the load information to determine whether the new cache is overloaded.

39. (previously presented) A computer program product as recited in claim 38, wherein load information is periodically received from the new cache system.

40. (currently amended) A computer program product as recited in claim 38, wherein the load information further indicates a number of buckets to shed or add, wherein the portion of the number of buckets that were previously shed from the new cache system, ~~the portion of the unassigned buckets,~~ and the portion of the assigned buckets are equal to the indicated number of buckets to shed or add.

41. (currently amended) An apparatus for assigning traffic buckets to a cache system, comprising:

Application No.: 09/608,549
CISCP151/3223

means for when a new cache system starts up in a cache cluster having a plurality of cache systems among which a plurality of total buckets are to be allocated, determining a full bucket allocation for the new cache system and assigning a portion of the full bucket allocation to the new cache system that was determined at start up;

means for periodically determining a load of the new cache system;

means for each time it is periodically determined that the new cache system is underloaded and buckets have not been previously shed from the new cache system, slowly assigning a portion of the full bucket allocation that was determined at startup unless the cache cluster is operating at a maximum load;

means for ~~when buckets have not been previously shed from the new cache system and~~ each time it is periodically determined that the new cache system is underloaded and buckets have been shed previously from the new cache system, slowly re-assigning ~~or~~ a portion of ~~previously shed the~~ buckets that have been previously shed from the new cache system to the new cache system ~~when buckets have been previously shed from the new cache system,~~ unless the cache cluster is operating at a maximum load; and

means for each time it is periodically determined that the new cache system is overloaded, shedding a portion of the buckets previously assigned to the new cache system,

wherein each bucket portion corresponds to a portion of the total traffic being handled by the cache cluster.